

UNIVERSIDAD DE CONCEPCIÓN  
DEPARTAMENTO DE PSICOLOGÍA  
ASIGNATURA TÉCNICAS DE EVALUACIÓN PSICOLÓGICA I

LA ESCALA DE INTELIGENCIA DE BINET-SIMON  
(ANTECEDENTES Y REVISIONES)  
Y  
LA ESCALA DE INTELIGENCIA DE F. GOODENOUGH  
(ANTECEDENTES Y MODO DE APLICACIÓN)

Apunte de circulación restringida para uso docente en la  
asignatura de Técnicas de Evaluación Psicológica I  
CONCEPCIÓN

## LA ESCALA METRICA DE LA INTELIGENCIA DE BINET-SIMON Y SUS REVISIONES

### ANTECEDENTES

Los primeros esfuerzos sistemáticos tendientes a buscar una solución a la carga que significaban los individuos con menor capacidad intelectual en el sistema educacional, se remontan al año 1894, en que la Municipalidad de París encargó al psicólogo Alfred Binet que buscara un método que permitiera seleccionar de entre los educandos a aquellos que no estuviesen capacitados para los estudios. Este hecho determinó no sólo el surgimiento de los test de Inteligencia en Francia, sino un modo de enfrentar el problema orientado exclusivamente por consideraciones prácticas. Lo que la Inteligencia (o su falta) significaba conceptualmente, no preocupó a Binet, llegando éste a afirmar que inteligencia era lo que medían sus pruebas.

Distinto fue el caso de Francis Galton, en Inglaterra. Llevado por su inquietud de estudiar los mecanismos de la herencia, especialmente de las capacidades intelectuales, desarrolló una serie de pruebas que publicó en 1893, tendientes a medir capacidades intelectuales, utilizando ya entonces la palabra "test" para designarlas.

Aun cuando la base de la psicología de los test se debe a Galton y sus seguidores, tuvieron mayor relevancia en el campo de la educación los trabajos de Binet. Sus primeros procedimientos para medir las capacidades de los niños en edad escolar los corrigió posteriormente con Th. Simon. Durante algún tiempo colaboró con ellos también Jean Piaget. Este, al darse cuenta mientras revisaba los ítem y las respuestas de los niños, que los errores que cometían los niños de diferente edad ante un mismo ítem eran distintos, pero que eran semejantes si los niños tenían una edad similar, comenzó a interesarse por el estudio de la evolución del razonamiento en los niños, tarea a la que dedicó el resto de su vida.

Una primera serie de pruebas ordenadas en función de su dificultad relativa, utilizada con el objeto de seleccionar a los alumnos que no tenían la capacidad intelectual necesaria para un aprendizaje escolar normal, fue publicada por Binet y Simon el año 1905. Comprendía 30 ítem o pruebas simples que no estaban agrupadas por edad.

La idea de la edad mental estaba implícita en esta primera versión, pero no fue formulada explícitamente hasta 1908, al publicarse la segunda versión, más elaborada.

## ELABORACION DEL TEST

La idea fundamental de estos test, según Binet y Simon, ha sido crear un gran número de pruebas fáciles de llevar a la práctica, realizables en poco tiempo, y dispuestas en orden de dificultad creciente; ensayar estos ítem en un gran número de niños de diferentes edades y anotar sus resultados; buscar las pruebas que sirven para una edad determinada, de modo que los niños más jóvenes sean incapaces de resolverlos y formar así una escala métrica de inteligencia que permita averiguar si un sujeto dado tiene el desarrollo intelectual correspondiente a su edad o está atrasado o adelantado y cuánto. Para establecer el grado de inteligencia propio de cada edad, se aplicó esta serie de pruebas a niños de distintas edades y se determinó qué proporción de ellos resolvía un determinado ítem. Si el 65% a 75% de los niños de una edad determinada resolvía bien el ítem, este será aceptado como un indicador de la inteligencia de esa edad. Se escogió este porcentaje ya que si lo resolvía menos que el 65% del ítem resultaba muy difícil y correspondía a una edad superior y si era resuelto por un porcentaje mayor que el 75%, resultaba demasiado fácil y correspondía a una edad menor. De aquí se derivó el concepto de Edad Mental (EM) que se basa en la consideración que a cada edad le corresponde un determinado grado de desarrollo intelectual o cognitivo. En este sentido, la EM indica meramente que la habilidad de un sujeto determinado, corresponde a la habilidad media de los niños de tal o cual edad

Considerando que a cada edad cronológica le corresponde en los niños normales una edad mental equivalente, propuso Binet los siguientes criterios diagnósticos:

- Si la EM resultaba 2 o más años menor que la EC, había retraso intelectual
- Si la EM resultaba 2 o más años superior a la EC, era adelantado o superior a lo normal
- Si la EM quedaba comprendida en un rango de  $\pm 2$  años la EC, el niño era normal

Estos criterios presentan algunos inconvenientes:

a) un mismo número de años de retraso tiene distintos significados según la edad del niño: un niño de EC= 4 años y EM= 2 años, tiene 2 años de retardo y ha logrado el 50% del desarrollo esperado, en tanto otro niño de EC= 8 años y EM= 6 años, tiene igualmente dos años de retardo (y por lo tanto la misma clasificación diagnóstica), pero ha logrado el 75% del desarrollo esperado a su edad.

b) Por otra parte, expresar el nivel de desarrollo intelectual como EM, induce a pensar que un niño de EC = 9 años y EM= 6 años (con tres años de retraso) tiene las mismas capacidades que otro niño de EC= 6 años y EM= 6 años (intelectualmente normal). Esto no es así pues la estructura de la inteligencia de un niño retrasado es diferente a la inteligencia de un niño normal.

c) El concepto de “retraso” implícito en este enfoque, induce a pensar que llegar a un determinado nivel de desarrollo cognitivo es sólo una cuestión de tiempo, por lo que un niño que presenta “retraso mental” podría llegar finalmente a niveles normales de desarrollo. Esto, en general, no es así.

Estos inconvenientes se resolvieron en las revisiones posteriores, con la introducción de cálculo de Coeficiente Intelectual (CI), propuesto por William Stern en 1912, y cuya fórmula es:

$$CI = \frac{EM * 100}{EC}$$

y en la cual ambas edades (mental y cronológica) se expresan en meses.

### REVISIONES DE BINET

Binet publicó una segunda versión de su prueba el año 1908. Esta estaba constituida por un total de 59 ítem (29 más que la versión de 1905), distribuidos diferencialmente en los grupos de edad. El rango de aplicación va de 3 a 15 años y considera 5 ítem por año de edad, con un valor de 2 meses 12 días cada uno ( o 1/5 de año); salvo para los 4 años, que hay 4 ítem con un valor de 3 meses cada uno.

Para la valoración de los resultados se comenzaba el cómputo a partir de la edad en que el niño resolvía todos los ítem. Por cada 5 o 6 ítem adicionales que aprobaba el niño, se le agregaba un año más a su EM. No se consideraban fracciones de año menores a un semestre y el resultado se expresaba como EM.

Puesto que el test fue diseñado para determinar si los niños estaban en condiciones de ir a una escuela común, presentaba muchos ítem cuya resolución depende del grado de escolaridad del examinado.

La experiencia acumulada con la aplicación de la escala de 1908 en Francia y otros países, llevó a una revisión y modificación parcial que fue publicada en 1911. En general se mantuvo la estructura de la versión de 1908, con 5 ítem para cada edad, con la excepción del año IV, que tiene 4 ítem. La diferencia radica en el criterio de selección de los ítem, que en este caso consistió en elegir aquellos que eran resueltos por el 50 a 75 % de los niños de cada edad.

Esta versión de 1911 fue considerada un recurso confiable y constituyó la base sobre la cual se realizaron las revisiones posteriores.

## **ADAPTACIONES DE GODDARD (1911) Y KUHLMANN (1912) EN EEUU**

Las escalas de Binet fueron introducidas en USA por Goddard, que realizó la primera adaptación y tipificación de la Escala de Binet - Simon de 1909, para niños norteamericanos.

Las diferencias culturales y educacionales determinaron la necesidad de reordenar algunos de los ítem en relación al grupo de edad a que estaban asignados, manteniendo el número de 5 ítem para cada edad a partir del año V.

Goddard trabajaba en una escuela para deficientes mentales en el estado de Nueva Jersey y trató de establecer algún criterio para el diagnóstico de la oligofrenia. Supuso que un retraso de dos años en niños menores de 9 años, y un retraso de tres años en niños mayores, indicaba una deficiencia mental probable.

Esta versión de Goddard se popularizó en USA y ocupó un lugar de privilegio hasta que apareció la primera revisión de la Universidad de Stanford, en 1916.

Kuhlmann publicó 3 revisiones de la prueba de Binet: una en 1912 (que no introduce grandes cambios sobre la versión de Binet de 1908, a excepción de extender la escala en ambos extremos), otra en 1922 y la última en 1939. Su aporte fundamental se refiere al énfasis que puso en que la situación de test fuera homogénea, es decir, estableció criterios para la aplicación del test.

## **REVISIONES DE TERMAN**

Lewis Terman, trabajando en la universidad de Stanford, adaptó el test de Binet – Simon para su utilización en EEUU, publicando su versión del test en el año 1916. En esta versión introdujo algunas modificaciones con respecto a la versión francesa: aumento el número de ítem a 90, extendiendo el rango de aplicación en un año más (hasta los 16 años), dejando 6 ítem para cada edad entre los 3 y 10 años, (de modo que cada uno equivale a dos meses), 8 ítem para el rango de 11 a 12 años (3 meses cada uno) y 6 ítem entre los 14 y 16 años (4 meses cada uno). Incorporó además la utilización del CI para expresar los resultados de la aplicación del test.

Los criterios que utilizó Terman para la elaboración de la escala (selección y ubicación de los ítem) contribuyeron efectivamente a señalar como inteligente a quienes lo parecían por su desempeño escolar o por su capacidad para resolver problemas.

Se preocupó, además, que el test, aparte de discriminar entre las diferentes edades, se comportara consistentemente en relación al puntaje total obtenido por grupos con distinta capacidad intelectual. El método de evaluación fue el mismo utilizado por Binet: se establece una EM base, equivalente al grupo de edad cuyos ítem han sido resueltos todos. Luego, por cada uno de los ítem adicionales resueltos, se le suman dos meses de EM (entre los años IV y X, ambos incluidos). La EM base, más la suma de los meses adicionales, constituye la edad mental del individuo. En consecuencia, ésta puede resultar hasta dos años mayor (18 años) que el límite de edad para la aplicación de la escala (16 años). De los 90 ítem que constituyen la prueba, 74 son regulares y 16 suplementarios. Una innovación importante la constituye la inclusión de la prueba de "vocabulario" y la posibilidad de aplicar una escala abreviada, utilizando en ese caso algunos ítem, señalados con un asterisco en la edición original. De acuerdo a los resultados obtenidos en la estandarización que realizó Terman en 1916, se obtuvo la siguiente distribución de los CI en la población:

**TABLA 1**

Rangos de CI y % de la población que comprende, según la estandarización de Terman del Test de Binet – Simon en 1916 en USA

RANGO DE CI	% DE SUJETOS
56 – 65	0,35
66 – 75	2,3
76 – 85	8,6
86 – 95	20,1
96 – 105	33,1
106 – 115	23,1
116 - 125	9,0
126 - 135	2,3
136 - 145	0,55

Terman propuso, además, una clasificación diagnóstica de los niveles de inteligencia y su relación con un conjunto de categorías que corresponden a los distintos grados de oligofrenia.

**TABLA 2**  
Rangos de CI y clasificación diagnóstica establecida por Terman

CI	CLASIFICACION
Sobre 140	Casi genio o genio
120-140	Inteligencia muy superior
110-120	Inteligencia Superior
90-110	Inteligencia normal o promedio
80-90	Torpeza raramente calificada como debilidad mental
70-80	Limítrofe entre simple torpeza y verdadera debilidad mental
80-70	Decididamente débil mental
50-70	Morón
25-50	Imbecilidad
Debajo de 25	Idiocia

### REVISIÓN STANFORD 1937

Esta nueva versión, elaborada por Terman y Merrill, consta de 129 ítem en cada una de sus dos formas (L y M). El objetivo es medir la inteligencia utilizando una escala tipo Binet. Al igual que la versión de 1916, se compone de pruebas destinadas a examinar la capacidad para comprender, advertir contradicciones, establecer semejanzas y diferencias, para reproducir dibujos, para repetir dígitos de memoria y muchas otras. A juicio de Terman y Merrill, estas pruebas permiten apreciar y estimar la aptitud para pensar abstractamente, o sea, la inteligencia.

La finalidad de la escala de 1937, es establecer el nivel de inteligencia sobre la base de un conjunto de aptitudes y funciones intelectuales que, en lo posible, no sufran la influencia de la educación.

En relación a la influencia de la educación, Terman y Merrill sostienen que la revisión de Stanford de 1916 tiene algunos ítem que están afectados por el grado de escolaridad. No debe dejar de considerarse los objetivos que se tuvieron al elaborar los test de Binet (diferenciar a los niños para ir a escuelas normales o diferenciales), aún cuando Binet y Simon al analizar sus propias escalas ya fueron intentando eliminar los ítem que demostraban ser demasiado sensibles a la escolaridad. Terman y Merrill intentaron en esta versión disminuir lo más posible el factor escolaridad, aunque no se puede decir que esté libre de él.

Esta escala es predominantemente verbal, lo cual corresponde al concepto de inteligencia de Terman, que la vincula a la capacidad de abstraer o pensar abstractamente.

Las escalas tipo Binet son muy sensibles a los factores culturales, no por su estructura, sino por sus contenidos. Esta revisión no escapa a estas influencias y según los autores, no lo intenta, ya que fue diseñada para el ámbito cultural norteamericano. En Argentina se usa la adaptación de Germaín Cebrain, un español que advierte de lo difícil de esta tarea y la versión Chilena corresponde al doctor Luis Tirapegui, quién destaca en su publicación que las pruebas de Terman fueron ensayadas durante años en el laboratorio de psicología experimental, para adaptarlas a las circunstancias especiales de nuestro medio ambiente.

La versión de Terman y Merrill de 1937 entrega los resultados en EM, con la cual se determina luego el CI. El trabajo de estandarización realizado por Terman y Merrill permite -en cierta medida- considerar este CI como un puntaje estándar derivado, que tendría una media = 100 y una DS = 16. En realidad, la media obtenida para las edades entre 21 y 18 años resultó, en ambas formas, ligeramente superior a 100, como también la DS real es algo mayor que 16:

Forma L	Media de medias = 104,00	Media de DS = 17,03
Forma M	Media de medias = 104,43	Media de DS = 16,75

La distribución que se obtuvo de los CI resultó ligeramente asimétrica, pero muy próxima a una curva de distribución normal, lo cual respalda la consideración de estos CI como puntaje estándar derivados y permite su comparación directa con los CI obtenidos utilizando otros test.

### **OTRAS REVISIONES DEL TEST DE BINET- SIMON.**

Zazzo y sus colaboradores quisieron adaptar mejor el test de Binet- Simon a los niños de su época y publicaron en 1949 una nueva versión, con algunas modificaciones técnicas que se referían especialmente a las consignas, a la revisión del orden de los ítem y a los baremos. Estaba destinada a ser aplicada entre los 3 y 14 años de edad. Años después, volvieron a tomar la escala de 1949 con el fin de mejorar su valor de test de desarrollo. El resultado de este trabajo fue publicado el año 1966 con el nombre de Nueva Escala Métrica de la Inteligencia (NEMI). Está destinada también a las edades entre 3 y 14 años y presenta varias diferencias con las versiones anteriores, tanto en sus cualidades como en su procedimiento de aplicación y análisis de los resultados.

De esta prueba existe una versión en español, publicada en 1960 por la editorial Kapelusz (Buenos Aires) y una estandarización para Chile que no alcanzó a concluirse. De ella se dieron a conocer en 1972, los resultados correspondientes a las edades de 7 a 16 años, en un apunte mimeografiado de la cátedra de psicología educacional de la Escuela de Psicología de la Universidad de Chile. Esta versión es poco conocida y pareciera no ser utilizada, a pesar de su adecuación a nuestra realidad y a sus buenas características psicométricas.

## LA ESCALA DE INTELIGENCIA DE F. GOODENOUGH

### ANTECEDENTES

La psicóloga Florence Goodenough es la creadora de un test de máxima simplicidad en sus fundamentos, administración y evaluación. Es una técnica que nos permite medir la inteligencia general por medio del análisis del dibujo de la figura humana.

Consiste en una única prueba de dibujar un hombre. La evaluación se reduce a computar el número de detalles acertados que incluye la figura realizada. Se finaliza convirtiendo ese puntaje en Edad Mental y luego en Coeficiente Intelectual.

### FUNDAMENTOS

¿Por qué el análisis del dibujo de la figura humana hace posible estimar la inteligencia del dibujante? ¿Cuál es el fundamento del test? La autora partió de una hipótesis de trabajo muy sencilla: cuando el niño traza la figura humana sobre un papel no dibuja lo que ve, sino lo que sabe al respecto y por lo tanto, no efectúa un trabajo estético, sino intelectual; ofrece una expresión no de su capacidad artística, sino de su repertorio conceptual. El volumen de ese repertorio conceptual, de este "saber" crece con la edad mental, y este proceso se refleja en el dibujo de la figura humana que hace el niño de 5 hasta por lo menos 10 años. La derivación psicométrica de este hecho es inmediata: la valoración cuantitativa del "saber" implicado en la representación humana típica de cada edad, suministra una norma para establecer, ante un dibujo particular, la edad mental de su dibujante.

¿Por qué la medición de este "saber" comporta la medición de la inteligencia? Según Goodenough al medir el valor de un dibujo se mide el valor de las funciones de asociación, observación, analítica, discriminación, memoria de detalles, sentido espacial, juicio, abstracción, coordinación viso-manual y adaptabilidad. En su sencillo acto de dibujar la figura de un hombre (de expresar su saber de la figura humana), el niño dibujante debe activar diversos recursos mentales: ASOCIAR los rasgos gráficos con el objeto real; ANALIZAR los componentes del objeto y REPRESENTAR, VALORAR y SELECCIONAR los elementos característicos; ANALIZAR las relaciones espaciales (posición); FORMULAR JUICIOS de relaciones cuantitativas (proporcionalidad); ABSTRAER, esto es, reducir y simplificar las partes del objeto en rasgos gráficos; y finalmente, COORDINAR su trabajo viso-manual y ADAPTAR el esquema gráfico a su concepto de objeto representado.

## LA ESCALA

Tras un estudio largo, complejo y esforzado de vastas colecciones gráficas de miles de niños de entre 2 y 15 años de diferente sexo, nacionalidad, razas, grado y rendimiento escolar, de elaboraciones estadísticas escrupulosas y reiteradas, así como rigurosas y limpias verificaciones de su validez y confiabilidad, Goodenough logró finalmente establecer su Escala y Baremo. Esa escala no es sino el inventario de los rasgos gráficos (ítemes) que mejor traducen el repertorio infantil del mencionado "saber" de la figura corporal. El catálogo permite comparar el grado de completación y perfección de un dibujo por la mera presencia o ausencia de ítemes, establecidos desde ciertos puntos de vista empíricos. Cabe distinguir estos 8:

- 1) Cantidad de detalles representados: ¿Se ha representado la cabeza? ¿las piernas? ¿el cabello? ¿el número correcto de dedos? etc.
- 2) Proporcionalidad: ¿El tronco es más largo que ancho? ¿La cabeza es proporcionada a los brazos y las piernas? etc.
- 3) Bidimensionalidad: ¿El cuello está representado en dos medidas o por una sola línea? ¿Y la boca y la nariz? etc.
- 4) Transparencia: Las prendas de vestir son opacas o transparentes? ¿Y los cabellos?
- 5) Congruencia: ¿Los miembros están unidos al tronco? ¿El cabello no excede el contorno de la cabeza? ¿Las diversas prendas de la vestimenta armonizan o son discordantes entre sí? etc.
- 6) Plasticidad: ¿El pulgar está en oposición? ¿La mano se diferencia del brazo y los dedos? ¿Las orejas están correctamente ubicadas?
- 7) Coordinación viso-motora: ¿El dibujo muestra seguridad de trazado, sin entrecruzamientos o sinuosidades? ¿Las facciones están simétricamente distribuidas?
- 8) Perfil: ¿Ha logrado dibujar el perfil? ¿Con cuántos errores?

Goodenough construyó su escala seleccionando 51 ítem (unidades de la escala) que se ajustan a uno o varios de esos criterios y los catalogó por orden de complejidad creciente. Luego computó el número de ítem (puntaje bruto), reunidos en una gran masa de niños, calculó el puntaje medio, (normas) de los niños normales de cada edad, tabuló esas normas y obtuvo la tabla (baremo) para determinar la edad mental de cualquier niño a partir de su puntaje bruto.

La ejecución en el test se evalúa verificando en el dibujo la presencia o ausencia de cada uno de los ítem de la escala acreditándole un punto por cada ítem cuyos requisitos se satisfagan, desde el primero (en que se otorga un punto por la mera presencia de la cabeza) hasta el último (en que se exige la representación de un perfil correcto).